UNCLASSIFIED

Defense Technical Information Center Compilation Part Notice

ADP021702

TITLE: Overview of Sensor Fusion Research at RDECOM NVESD & Recent Results on Vehicle Detection Using Multiple Sensor Nodes

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: Proceedings of the International Conference on International Fusion [6th]. Held in Cairns, Queensland, Australia on 8-11 July 2003. Volume 1: FUSION 2003

To order the complete compilation report, use: ADA442007

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP021634 thru ADP021736

UNCLASSIFIED

Overview of Sensor Fusion Research at RDECOM NVESD & Recent Results on Vehicle Detection Using Multiple Sensor Nodes

Philip Perconti
James Hilger
Science & Technology Division
RDECOM/CERDEC NVESD
Fort Belvoir, VA, U.S.A.

Murray Loew
Department of Electrical and Computer Engineering
The George Washington University
Washington, DC, U.S.A.

Abstract - *The US Army RDECOM CERDEC Night Vision & Electronic Sensors Directorate has a dynamic applied research program in sensor fusion for a wide variety of defense & defense related applications. This paper provides an overview of the on-going research at NVESD related to fusing a mixture of active and passive sensors for countermine, dismounted & mounted soldiers, aviation and unattended ground sensor applications. Highlighted are new techniques in image registration and sensor fusion that enable the detection of moving vehicles with a network of image and acoustic sensors. A set of experiments was designed and conducted using a variety of vehicles and scenarios. The incremental value (in terms of error probability) added to the acoustic information by the image sensors (visible- and infrared-wavelength cameras) is assessed in combination with the fusion techniques themselves. The approach specifically accounts for the effects of location, speed, weather, and background (acoustic, visible, and infrared). Sensor fusion for detection and classification is performed at both the sensor level and the feature level, providing a basis for making tradeoffs between performance desired and resources required. Several classifier types are examined (parametric, nonparametric, learning). The combination of their decisions is used to make the final decision.*

Keywords: Tracking, fusion, sensor fusion, vehicle detection, detection, location.

1 Introduction

There are multiple areas of applied sensor fusion research at the US Army RDECOM CERDEC Night Vision & Electronic Sensors Directorate (NVESD). These areas include internal efforts, collaborative efforts with universities, as well as contract efforts with private industry. Many of the early efforts at NVESD during the 1980's and 1990's showed that multiple collocated sensors could be used to improve both human and sensor system performance in detecting targets and other objects of interest. These early efforts fused image data with non-image data in order to improve the performance of algorithms processing the image. Current efforts use multiple, more robust types of sensor data and fusion techniques in order to improve overall system performance as measured by Probability of Detection (P(d)) and False Alarm Rate (FAR). All of these efforts highlight the trend within the Department of Defense (DoD) for using multiple sensors, either collocated or dispersed, to detect objects of interest. This trend has fueled efforts such as those focused on mine and minefield detection (countermine), mounted and dismounted soldiers, and unattended ground systems.

Current countermine research has shown that no single sensor technology will likely be able to address the requirements associated with the detection of mines or mine fields [1]. There are efforts within countermine to fuse Ground Penetrating Radar (GPR) and acoustical data [2], GPR and seismic and acoustical data [3], and image data and acoustical data [4]. These efforts utilize both multi-sensor and multi-look fusion to increase performance over that achieved through the use of single sensors.

Multiple approaches to sensor fusion are being utilized in research for dismounted and mounted soldier systems [5, 6]. Early work focused on fusing long wave IR, visible and near IR out to 900 nm. This work was aimed at fusing the complementary information found in thermal and low light level sensors for such applications as helicopter pilotage and night driving [7, 8]. Other research includes the utilization of imaging sensors to simultaneously collect data from multiple long wave IR spectral bands for target detection. Sensor data are fused at multiple levels in order to maintain or increase the P(d) while reducing FAR through clutter rejection. One effort is further partitioning the IR bands in order to improve overall system performance in detecting objects of interest under various conditions. These types of systems usually have multiple imaging sensors or/and non-imaging sensors collocated on the same platform. Related sensor efforts are focused on the development of focal planes

sensitive to multiple IR bands and adaptive hyperspectral focal planes that adapt based on spectral information. [9]

Some of the approaches include the utilization of sensor data from other non-collocated sensors. Images, image chips, and/or informational data are utilized in the sensor fusion. These other sensors can be located on unmanned air vehicles and unattended ground systems. The idea is to get multiple looks of a field of regard from multiple sensors in order to detect objects of interest.

Unattended Ground Systems (UGS) are attractive in the sense that they can be deployed in an area for surveillance purposes over long periods of time. These sensors consist of sensor nodes with imaging (IR, visible) and non-imaging (seismic, acoustic, magnetic) sensors and communication capability to establish sensor networks with each other as well as command and control facilities. When UGS are deployed, they self configure into a network to perform surveillance to improve situational awareness. Collaboration between UGS allows for multi-sensor, multi-look data to be passively collected and fused such that target characterization can occur either at the node, or at a remote location. Research at NVESD is focused on demonstrating detection algorithms and related node networking schemes as well as studying long-term deployment impacts for these sensor systems. One research program in this area at NVESD is the Sensor Fusion Testbed .

2 Sensor Fusion Testbed

The Sensor Fusion Testbed (SFTB) is an environment consisting of three outdoor nodes and an indoor base-station and associated software to support the development and testing of imaging and non-imaging sensors and algorithms. Its purpose is to provide a consistent and well-defined testing facility for advanced sensors and associated detection and fusion algorithms. The SFTB will evolve in the sense that it will be augmented with additional nodes of varying capability and algorithms over time. Through the SFTB, analysis of the impact to advanced sensors and associated algorithms deployed over long periods of time on station in urban and rural environments can be accomplished.

2.1 SFTB nodes and base-station

The SFTB nodes have imaging and non-imaging sensors and associated computer hardware. Two of the nodes have advanced uncooled Infrared (IR) cameras. The IR cameras have 40 degrees horizontal by 30 degrees vertical field of view and output 160 by 120 pixel images where each pixel is 12 bits in length. The other node has a color camera with 52.06 degrees horizontal and 30.45 degrees vertical field of view. The full resolution 640 by 480 pixel digital output can be down sampled to produce a 160 by 120 pixel image where each pixel is 24 bits (8 bits red, 8bits green, 8bits blue). The IR and color cameras are attached to 2-axis, pan and tilt, gimbals controlled by software. Each node has an acoustical array consisting of 7 microphones arranged such that 6 microphones are equally spaced along the periphery of a 4-foot diameter circle and one microphone is in the center of the array. Figure 1 shows a picture of one of the nodes. The camera and gimbal sit atop a tripod with the acoustical array mounted at the base of it as shown in the figure. Each node has a Global Positioning System (GPS) and a personal computer, ruggedized for outdoor use, to hold the necessary interface electronics as well as provide a processor for those applications that want to do processing of the sensor data at the node. The personal computer also has data storage to record both sensor input data as well as any algorithm processing results. Cabling shown in Figure 1 are sensor interface cables to the personal computer in order to allow it to be up to 50 feet away from the sensor nodes.

Figure 1. Sensor fusion testbed node.

The SFTB nodes are connected to a base-station through a wireless Ethernet link. The base-station is a personal computer inside a building that provides the SFTB with a number of capabilities and services. A user interface has been written to allow each SFTB node to be controlled from the base-station. Sensor outputs can be transmitted to the base-station to provide another level of sensor fusion capability for those applications that require it. With sensor data processing capability at each node as well as at the base-station, multiple kinds of sensor fusion applications can be setup for analysis and demonstration. Weather data is obtained from a weather station whose output weather measurements are input to the base-station.

2.2 SFTB deployment

The SFTB is deployed initially in a triangular location illustrated in Figure 2. The nodes are labeled 1 through 3 with the imaging sensor type in parenthesis. Again, all nodes have a 7 element acoustical array. The order of the nodes can be switched around for testing purposes depending on the particular application. The triangular location represents a complex area with various backgrounds, road directions, building structures and noise sources. The surrounding area has numerous continuous and periodic man made noise sources, various buildings at different distances, aspect angles, shapes and construction materials, various traffic patterns and road surfaces, various vegetation, and wildlife. Due to the road structure around the triangular location, various distances and directions to objects of interest can be setup for testing purposes. The nearby parking lots provide additional opportunities for testing.

While the triangular location was chosen in part due to the complexity offered from a myriad of noise sources, types of clutter, and movement of objects of interest, the SFTB is also portable. Portability allows the nodes and base-station to be taken to other locations for more focused data collections, detection and fusion algorithm tests and/or analysis. This enables other types of environmental conditions from more rural to more urban compared to the triangular location to be included in any analysis.

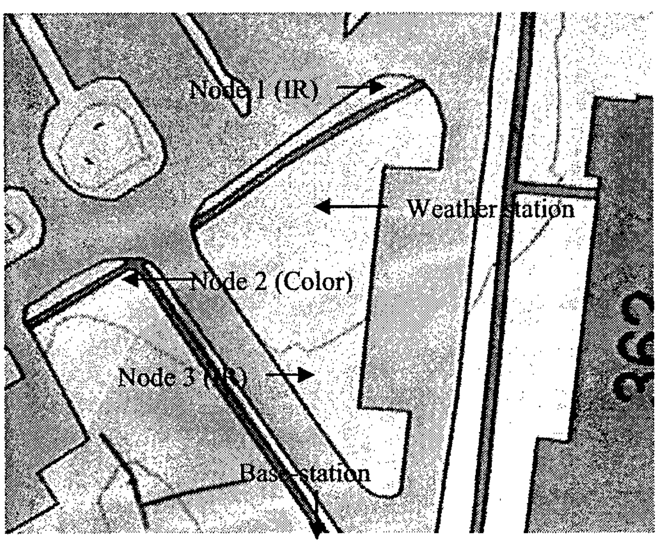

Figure 2. Sensor fusion testbed deployment.

2.3 SFTB vehicle detection demonstration

A vehicle detection demonstration has been implemented into the SFTB during its development and integration. This simple demonstration was implemented to ensure components within the SFTB were functioning properly, demonstrate the utility of the SFTB in such a system application, serve as a baseline application to guide algorithm enhancements and experiments, as well as guide the integration of the SFTB into the NVESD algorithm evaluation facility. The vehicle detection demonstration exercises all nodes within the SFTB as well as the base-station. This demonstration serves as the baseline for which to compare other sensor fusion algorithm and node concepts.

The vehicle detection demonstration uses outputs from the processing of the acoustical data to cue the imaging sensors to point in a particular direction and snap a picture. Figure 3 illustrates the concept. The power spectrum of the acoustical signal from one acoustical array is shown in the top of the figure. The snapshots are those of a light truck passing the node. The truck approaches the node cueing the imaging sensor to point in the direction of arrival. The imaging sensor follows the truck noise as it passes, continuously snapping pictures based on the processing of acoustical data. It should be noted that Figure 3 shows the outputs of only a single node to illustrate the concept. Outputs from the others nodes are available also and are used in the vehicle demonstration.

To accomplish this concept, each of the three sensor nodes within the SFTB has been programmed with a version of the MUSIC algorithm that has been adapted by the U. S. Army Research Laboratory to determine the angle of arrival for vehicles of interest [10,11]. Acoustical data at each node is processed using this MUSIC algorithm. The resultant angle of arrival from each node is transmitted via wireless Ethernet to the base-station. The base-station then charts the angles of arrival to determine if any intersect. If two or more angles intersect, then a detection is declared. The angles of arrival are displayed in a graphical user's interface for the operator.

2.4 SFTB data analysis

The SFTB was designed to support development and testing of imaging and non-imaging detection and fusion algorithms. This also includes collection of data to do offline algorithm development and evaluation as well as data analysis. Algorithms and demonstrations developed and implemented on the SFTB support the focus of studying the phenomenology of an area's environmental impact to sensors and associated detection and fusion algorithms. This addresses the shortfall of knowledge, and supporting data, of how well detection and fusion algorithms perform when deployed over long periods of time. Since the driver for sensor fusion is increased system performance in terms of false alarm rates and detection probabilities, through the SFTB these metrics can be analyzed in terms of variables such as road types,

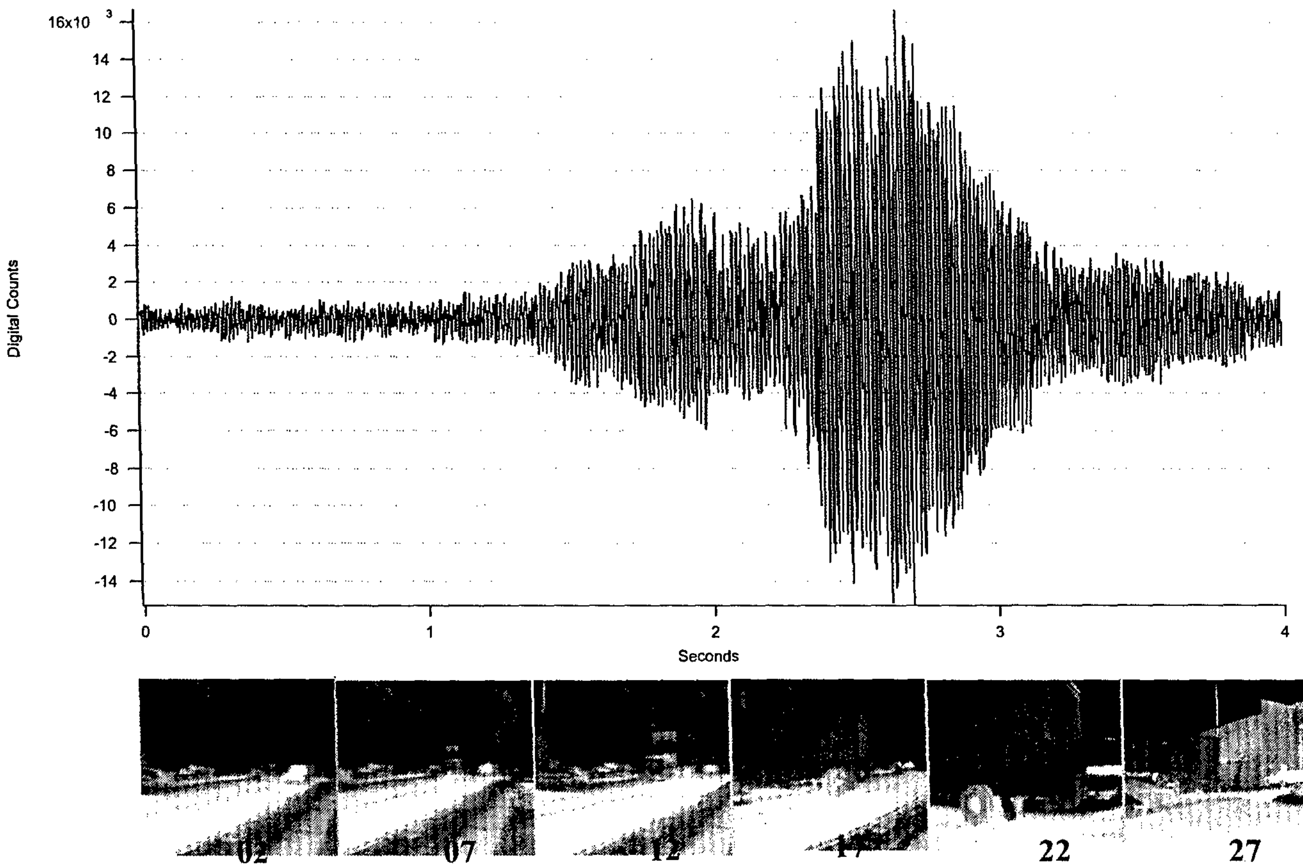

Figure 3. Sensor fusion testbed demonstration program outputs.

environmental conditions, random natural and man made noise sources, sensor data characteristics over seasonal changes, as well as the resulting physical system level attributes such as average power consumed.

The volumes of multi sensor data generated by long-term studies using the SFTB requires methods and supporting automation to truth, catalog, and database in real time. The current Aided Target Recognition (ATR) algorithm evaluation facility at NVESD is setup to post process sensor data to generate statistical information after the data have been collected in the field, batch process ATR algorithm evaluations, and enable analysis of data and algorithm performance. Many man-years of effort have established methods and supporting tools and utilities to automate this facility such that it is a premier facility within the U. S. Army. Researchers are able to query Oracle databases containing numerous metrics calculated on sensor data such as mean background pixel intensities, mean target pixel intensity, total number of edge pixels, etc. to study various aspects of targets and backgrounds. Meteorological data and individual target data such as vehicle surface temperature are also available in the database. This facility is being enhanced to enable multi sensor data obtained through the use of the SFTB, as well as all algorithm outputs, both single sensor and fused, to be ground truthed, cataloged, and analyzed continuously in real time.

Included in the analysis will be anomaly detectors operating in real time. Over long periods of time, it is assumed that much of the data obtained from one location will settle to some activity level as measured by a particular set of statistics or algorithm output metrics such as vehicles detections or false alarms per hour. The activity level may slowly drift over time due to multiple factors, but will be fairly stable. Once a norm is learned, deviations from this norm can be highlighted for immediate further in-depth analysis. Overall system level performance can and will be analyzed continuously, but deviations are important enough to signal their occurrence. The NVESD ATR algorithm evaluation facility is being augmented to support continuous real-time analysis as well as signaling deviations from the norm.

The analysis will focus on a "top 10" list of variables. Objects of interest are vehicles from 4 classes: car, pickup, SUV, and light truck as well as human. Each object is in a state as determined by tire type, tire air pressure, and general overall condition. Surface types of interest include asphalt, concrete, gravel, dirt, and grass. The objects of interest could be operating near noise

sources that are man made or naturally occurring. Weather includes wind, temperature, and precipitation during the four seasons. All of this would be monitored using sensor nodes in any number of possible arrangements. Each sensor node could be composed of sensors of differing capabilities such as seismic, acoustic, imaging IR, proximity IR, and visible, both monochrome and color. Each node could be collaborating with its neighbor, groups of neighbors, or command station. This leads to what data to fuse, and how to fuse it such that a minimal number of sensor nodes are required to effectively cover an area under surveillance. Also of interest is the possibility of learning the background noise, both periodic and non-periodic, over extended periods of time in order to eliminate information not of interest.

Whenever sensor nodes are deployed in the field, the variables listed above are those that impact overall system performance. Understanding the impacts over long periods of time will help improve overall sensor nodes and their capability. Collecting data to aid experiments addressing these variables will be accomplished through data collections using the SFTB and the NVESD algorithm evaluation facility.

3 Sensor Fusion Data Collection

There are 2 distinct purposes for collecting multi sensor data using the SFTB. The first purpose is the generation of an initial problem set having simultaneously collected acoustical and image sensor data for work in multi sensor fusion for vehicle detection. This will provide data for initial research and developments addressing the "top 10" list of variables from the previous section. Problem set data is typically divided into parts with one part being utilized in the development and training of algorithms and the other part sequestered for algorithm evaluation. The second purpose of this data collection is to aid in the establishment of the necessary methodologies and tools to collect, catalog, and ground truth simultaneously collected multi-sensor data. As described in the previous section, the NVESD algorithm evaluation facility is being enhanced to ground truth, catalog, and analyze continuous, real-time, multi sensor data and algorithm outputs. This initial data collection is referred to as the SFTB Local Site (SFTBLS) data collection and it will be repeated over the year such that seasonal conditions are captured.

3.1 SFTB local site data collection

The SFTBLS data collection will generate a problem set containing what is referred to as signature data. These data will be collected in a controlled fashion with known vehicles moving at known predetermined speeds operating on a predetermined course. The SFTB sensor nodes will be setup at known locations and surveyed in such that pointing angles and distances between nodes are known. The imaging sensors will be pointed in a fixed direction that will allow for vehicles to enter and leave the sensor node's field of view. Data will be collected in 3-minute segments to allow time for vehicles to traverse the predetermined course at the slowest rate and keep the data in manageable segments for fusion algorithm development purposes. Each vehicle will traverse the predetermined course at 5, 10, 15, 20 and 25 miles per hour. The course has parallel asphalt and gravel road surfaces. Data will be collected with multiple vehicles traversing the course at one time. All vehicle positions and movements will be recorded through the use of Global Positioning System (GPS) units. Acoustical data will be collected simultaneously with the image data. The acoustical data recorded will be that from the outputs of the 7 microphones after digitalization.

Sensor data will be collected on each of the vehicles as they are rotated about their center axis at 15-degree intervals at a fixed distance from a sensor node with a visible camera and a sensor node with an IR camera. Acoustical data will also be collected from each vehicle as it is rotated and as it starts and revs its engine up to 2000 RPM. 2000 RPM was selected as the peak since it is equivalent to highway speed. Turntable data is important in that it highlights a vehicle itself rather than the vehicle in its surroundings, so that pristine vehicle signatures can be obtained.

3.2 SFTB future data collections

The SFTB will be utilized for future data collections to generate other problem sets whose data further addresses the "top 10" list of variables. Future collections will occur at the same location as the SFTBLS and else where. Along with location additional sensors, sensor nodes, and vehicles from the 4 classes will be utilized. These controlled, multi-sensor problem sets will be such that everything about the data will be known and recorded.

3.3 SFTB data availability

The SFTBLS data collection will generate an initial set of data that can be provided to others for use in sensor fusion efforts. At the completion of the data collection, the data will be checked and ground truthed. A portion will be sequestered for evaluation purposes. The rest will be made available to interested parties working in the sensor fusion community. As additional data collections are accomplished, those data will also be made available.

The data collected continuously, 24 x 7 using the SFTB setup in the triangle location will also be made available. This data can be obtained in segments lengths other than 3 minutes. Vehicles within this data will not be instrumented to the extent of those in the controlled data collections nor will turntable data be available for any of

the vehicles. The intention of this type of data is the long term analysis described earlier.

4 Sensor Fusion Data Analysis

As indicated above, acoustic data will cue and then guide the IR and visible sensors, which will acquire imagery, extract features, and decide whether this is a detection or false alarm. Multiple acquisitions over time for a given stimulus will be made and multiple classifiers used to fuse features from each of the sensors and make a decision. A set of training data will be analyzed to identify the useful features and be used for the design of the fusion algorithm. The techniques will be evaluated using jackknife procedures and also a separate test data set, as available.

4.1 Fusion errors

Image features are intended to describe points, lines, areas, and their temporal behavior. The ability to extract some of those features (e.g., texture) is dependent on resolution (and hence range). Postulation of various minimum target dimensions will enable bounds to be established for the extraction of resolution-dependent features. The maximum range for IR and visible measurements will depend on the specifications of the system. Sensor separation (baseline) will be determined in part by sensor resolution to ensure that the precision provided by each is comparable.

It is possible to compute the maximum location error when the fields of view of the sensors in an array are known. For example, for the typical SFTB geometry shown in Figure 4 (an isosceles triangle with two 100-unit legs) and for an assumed sensor field-of-view of 1.2 mrad,

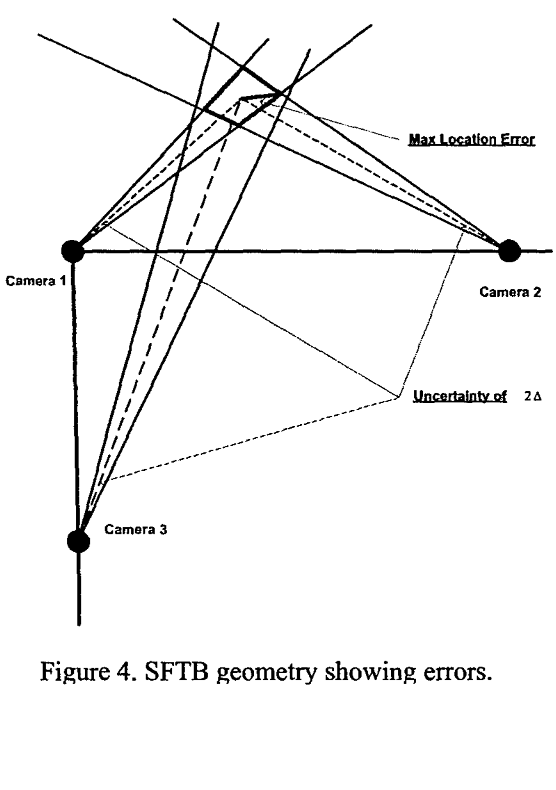

Figure 4. SFTB geometry showing errors.

the contours of iso-error shown in Figure 5 result. The three axes of the plot use the same units.

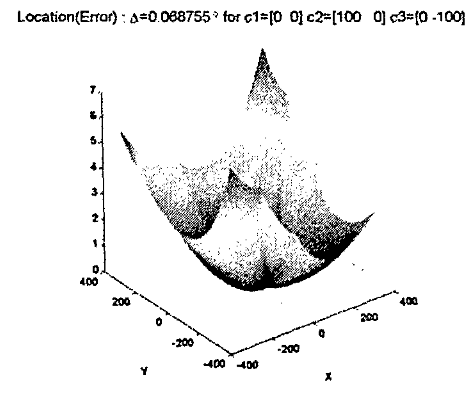

Figure 5. Contours of iso-errors for SFTB.

4.2 Fusion approaches

As stated earlier, all data collected with the SFTB will have associated ground truth in order to determine the image truth for vehicle location. Features can be extracted from the time-synchronized source data pertaining to vehicles and backgrounds.

The background will be learned by the system by periodic scanning of the environment throughout the pan/tilt range of each of the cameras at times during which targets are known to be absent; images acquired will be indexed by position and time and stored in a reference database. Background acoustic information will be acquired in the same way. The background data will be used to normalize the dynamic input data by subtraction and filtering.

Once the SFTB data and the turntable reference data are collected, classification will be performed using component classifiers to ensure expertise in various regions of the feature space. The use of clustering algorithms and problem-specific knowledge will allow the overall problem to be divided into regions, each of which will have its own classifier designed and tested. Several classifier types will be considered (parametric, nonparametric, learning). The combination of their decisions will be used to make the final decision. Initially, there will be fusion of features extracted at each array. The marginal contribution of the imaging sensors to performance will be estimated. Receiver operating characteristic (ROC) curves will be constructed.

Experiments will be performed to determine the estimated bounds on performance, as expressed by the ROC imposed by vehicle type, vehicle location, speed, and direction; lighting; weather (wind, rain, temperature); background (image and sound); and sensor positions.

These experiments will be expanded through the addition of nodes of various capabilities to the SFTB. The idea behind these experiments is to understand the operation of these types of sensors in an operational setting over long periods of time. The bounds discovered through this analysis will help guide the next generation of technologies for vehicle detection through the use of sensor fusion.

5 Conclusions

This article highlighted sensor fusion research efforts at the US Army RDECOM CERDEC Night Vision & Electronic Sensors Directorate. The research is broad in the sense that it includes activities from the development of advanced sensors, sensor nodes, fusion algorithms for detection to demonstration of these systems in the field, under real world conditions and situations. The sensor fusion testbed is an example of an effort to understand the long-term impact of deployed UGS as well as serve as a testing facility for advanced sensors and associated detection and fusion algorithms.

References

[1] Maksymonko, G. B. et. al., "Robust detection and fusion of mine images Proceedings of SPIE - the International Society for Optical Engineering, vol. 4742, 2002, pp. 142-149. USA

[2] Witten, T. R. et. al., "Fusion of ground penetrating radar and acoustics data", Proceedings of SPIE - the International Society for Optical Engineering, vol. 4742, 2002, pp. 903-910. USA

[3] Bradley, M. R. et. al., "Fusion of acoustic/seismic and ground-penetrating radar sensors for antitank mine detection", Proceedings of SPIE - the International Society for Optical Engineering, vol. 4394, 2001, pp. 979-990. USA

[4] Williams, A. C. et. al., "Development of a robust algorithm for detection of mine targets in image data from electro-optic and acoustic sensors", Proceedings of SPIE - the International Society for Optical Engineering, vol. 4394, 2001, pp. 943-951. USA

[5] Reese CE, Bender EJ, "Advancements of the head-tracked vision system (HTVS)", Proceedings of SPIE - the International Society for Optical Engineering, vol. 4711, 2002, pp.105-16. USA.

[6] Reese CE, Bender EJ, "Multispectral image-fused head-tracked vision system (HTVS) for driving applications", Proceedings of SPIE - the International Society for Optical Engineering, vol.4361, 2001, pp.1-11. USA.

[7] Steele, PM., Perconti, P., "Part task investigation of multispectral image fusion using gray scale and synthetic color night vision sensor imagery for helicopter pilotage.", Proceedings of SPIE - the International Society for Optical Engineering, vol. 3062, 1997, pp.88-100. USA.

[8] Perconti, P., "Noise limited resolution of the Advanced Helicopter Pilotage helmet mounted display and image intensified camera.", Proceedings of SPIE - the International Society for Optical Engineering, vol. 3058, 1997, pp. 254-63. USA.

[9] Horn, S., et. al., "Fused reflected/emitted light sensors", Proceedings of Spie - the International Society for Optical Engineering, vol. 4369, 2001, pp.1-13. USA.

[10] T. Pham and B. Sadler, "Aeroacoustic wideband array processing for detection and tracking of ground vehicles", Journal of the Acoustical Society of America, vol. 98, no. 5, pt. 2, p 2969, 1995.

[11] T. Pham and M Fong, " Real-time implementation of MUSIC for wideband acoustic detection and tracking" SPIE AeroSense 97: Automatic Target Recognition VII, April 1997.